Perceptual learning alters post-sensory processing in human decision making

Jessica A. Diaz[1], Filippo Queirazza[1] and Marios G. Philiastides[1]*

Affiliations:

1. Institute of Neuroscience & Psychology, University of Glasgow,

58 Hillhead Street, G12 8QB, Glasgow, UK

*Correspondence to: Marios.Philiastides@glasgow.ac.uk

Abstract

An emerging view in perceptual learning is that improvements in perceptual sensitivity are not only due to enhancements in early sensory representations but also due to changes in post-sensory decision processing. In humans, however, direct neurobiological evidence of the latter account remains scarce. Here, we trained participants on a visual categorization task over three days and used multivariate pattern analysis of the electroencephalogram to identify two temporally-specific components encoding *sensory* (Early) and *decision* (Late) evidence, respectively. Importantly, the single-trial amplitudes of the Late, but not the Early component, were amplified in the course of training and these enhancements predicted the behavioural improvements on the task. Correspondingly, we modelled these improvements with a reinforcement learning mechanism, using a reward prediction error signal to strengthen the readout of sensory evidence used for the decision. We validated this mechanism through a robust association between the model's decision variables and our Late component's amplitudes indexing decision evidence.

Introduction

Consider an image intelligence analyst inspecting a large array of noisy CCTV or satellite images in order to identify targets that might pose a real security threat. Her ability to perform this task successfully depends on her years of experience in interpreting such images. This example highlights that training and experience are required to induce long-lasting improvements in our ability to make decisions based on ambiguous sensory information a phenomenon commonly referred to as perceptual learning [1,2]. Despite the prevalence and obvious utility of this phenomenon in everyday life (e.g. learning in an ever-changing environment to make better predictions and plan future actions), its neural substrates and how these affect decision-making remain elusive.

Several psychophysical studies offered evidence linking perceptual learning with enhancements in early sensory representations [3-9] and with changes in post-sensory processing relating to attention and decision making [10-12]. In line with the latter account (i.e. late influences), recent experimental work in non-human primates (NHP) [13,14] offered compelling evidence that perceptual learning in decision making can affect how early sensory representations are interpreted downstream by higher-level areas to form a decision.

Correspondingly, recent functional magnetic resonance imaging (fMRI) experiments in humans started to address the question of whether perceptual learning affects later processing stages [15-18]. To date, however, little has been done to exploit *time-resolved* electrophysiological signatures that can accurately differentiate between early stimulus encoding and late decision-related processing. Here, we test the extent to which perceptual learning alters post-sensory encoding of decision evidence in humans by

recording electroencephalography (EEG) data during a face/car perceptual discrimination experiment (Fig. 1a) over the course of three days. Previously, using this task and single-trial multivariate discriminant analysis of the EEG we identified two temporally distinct neuronal components that discriminated between the stimulus categories: an Early component that occurred around 170 ms after stimulus presentation and a Late component that occurred around 300 ms post-stimulus [19-23].

We showed that compared to the Early one, the Late component was a better predictor of behaviour [20], it systematically shifted later in time with perceived task difficulty [19] and it was a significantly better predictor of trial-by-trial changes in the rate of evidence accumulation (i.e. drift rate) in a drift diffusion model [19,23]. Finally, while the Early component amplitudes remained unaffected when the same (face/car) stimuli were coloured red or green and the task was switched to colour discrimination those of the Late component were reduced almost to zero [19,22,23]. Taken together these findings indicated that the Early component encodes the incoming sensory evidence, whereas the Late component indexes, post-sensory, decision-relevant evidence. These previous findings are intriguing because they establish a benchmark against which to evaluate the extent to which perceptual learning influences earlier vs. later stages of decision making.

Specifically, here we test how activity associated with each of these Early and Late EEG components is affected by training. We hypothesize that if perceptual learning primarily alters post-sensory encoding of decision evidence, discrimination performance for our Late but not the Early component should systematically increase across the three training sessions. Similarly, as perceptual sensitivity improves with training we expect the Late component to move earlier in time, reflecting a decrease in perceived task difficulty. Moreover, our ability to exploit single-trial variability in the EEG will offer a

mechanistic characterization of these effects by establishing whether improvements in discrimination are a result of gain modulation (i.e. amplification of the differential response) of the component amplitudes, a reduction in the trial-to-trial variability (i.e. noise) of the component amplitudes or both.

Finally, we explore the possibility that these improvements can be understood in terms of a reinforcement learning (RL) mechanism [14,17,24-26], whereby the connections between early and late decision processing stages are strengthened via a reward prediction error, gradually enhancing the readout of relevant information and leading to improved perceptual sensitivity.

Results

We collected behavioural and EEG data from 14 participants during a speeded face vs. car categorization task using noisy stimuli that varied in the amount of available sensory evidence (i.e. phase coherence of the stimuli). Visual feedback was provided for each response prior to the presentation of the next stimulus (Fig. 1a). Participants performed the same task on three consecutive days. Using a mixed-effects logistic regression analysis, we found that accuracy was significantly improved ($x^2_{df=1}$= 19.37, p < 0.001, Fig. 1b) over the three training days. Using a mixed-effects linear regression analysis, we found that reaction times (RT) were significantly reduced over the three training days ($x^2_{df=1}$ = 8.92, p < 0.003, Fig. 1c). We note that, as expected, we also found a main effect of stimulus difficulty, with accuracy increasing ($x^2_{df=1}$= 28.08, p < 0.001) and RT decreasing ($x^2_{df=1}$= 21.24, p < 0.001) with the amount of sensory evidence, respectively. There was no interaction between the amount of sensory evidence and training day on either measure (accuracy: $x^2_{df=1}$= 0.16, p = 0.68, RT: $x^2_{df=1}$= 0.383, p = 0.54).

Next, we sought to identify the Early (sensory) and Late (decision-related) EEG components that discriminate between face and car trials and investigate how these are affected by training. To this end, we used a single-trial multivariate discriminant analysis [27,28] to identify linear spatial weightings of the EEG sensors, which best discriminated between the two trial types. For each participant, we estimated, within short pre-defined time windows of interest, a projection in the multidimensional EEG space (i.e. a spatial filter) that maximally discriminated between the two categories on stimulus-locked data (Eq. 1; see Methods). Applying this spatial filter to single-trial data produced a measurement of the resultant discriminating component amplitude (henceforth y). Component amplitudes can be thought of as indexing the quality of the evidence in each trial, in that a high positive amplitude reflects an easy face trial, an amplitude near zero reflects a difficult trial, and a high negative amplitude reflects an easy car trial (Fig. 2a). We used the area under a receiver operating characteristic curve (i.e. Az-value) with a leave-one-out trial cross validation procedure to quantify the discriminator's performance (i.e. the degree of separation in the single-trial amplitude distributions associated with each stimulus category).

Our discriminator's performance as a function of stimulus-locked time revealed the presence of two temporally specific components (Fig. 2b; Early, mean peak time: 187 ms; Late, mean peak time: 431 ms), consistent with our previous work [19-23]. Most crucially, even though both the Early and Late components reliably discriminated between image categories, only the discrimination performance for our Late component appeared to systematically increase across the three training days. To formally test for this effect we extracted subject-specific peak Az-values for each of the Early and Late components and run a mixed-effects linear regression analysis with training day,

component (i.e. Early vs Late) and their interaction as separate predictors. We found a significant main effect of training day ($x^2_{df=1}$ = 7.61, p = 0.006), a main effect of component ($x^2_{df=1}$ = 5.0371, p = 0.025) and a significant interaction between the two ($x^2_{df=1}$ = 7.46, p = 0.006), indicating that discriminator performance for the Late component increased systematically across training days, whereas that of the Early component remained unchanged (Fig. 2c). Taken together, these results provide compelling evidence that it is primarily the encoding of the decision evidence in the Late component, rather than the sensory evidence in the Early component, that is being enhanced in the course of training.

In previous work [20], we showed that, unlike the Early component, the peak time of the Late component moved later in time as perceived task difficulty increased, consistent with longer integration times for more difficult decisions [29-31]. Here, we exploit this finding to provide additional evidence linking the Late component with the process of learning. Specifically, we hypothesized that the latency of the Late component should move earlier in time as learning unfolds (i.e. as choices become easier). Using a separate mixed-effects linear regression analysis we found a significant main effect of training day ($x^2_{df=1}$= 21.56, p < 0.001), a main effect of component ($x^2_{df=1}$= 51.3, p < 0.001;) and a significant interaction of the two ($x^2_{df=1}$= 51.75, p < 0.001) on component peak times, indicating that the Late component peak times were reduced systematically across training days, whereas those of the Early component remained unchanged (Fig. 2d). These findings reinforce the notion that it is the temporal dynamics of the Late decision-related component that change as a function of training.

To better understand the mechanism by which improvements in discrimination

157 performance for the Late component came about, we capitalized on the single-trial
158 variability in the component amplitudes. Specifically, we tested whether there was an
159 increase in the distance between the mean face and car component amplitudes in the
160 Late component ($\bar{y}_f - \bar{y}_c$; Fig. 3a), a reduction in the trial-by-trial variability around those
161 means ($\sigma(y_{c,f})$; Fig. 3b) or a combination of both. We ran a mixed-effects linear
162 regression analysis, with the amount of sensory evidence, training days, and their
163 interaction as separate predictors.
164
165 As expected from previous findings [20,22,23] we found a main effect of the amount of
166 sensory evidence on the means ($x^2_{df=1}$= 11.52, p < 0.001, Fig. 3c) but not on the
167 variance of these component amplitudes ($x^2_{df=1}$= 0.38, p = 0.53, Fig. 3d). Crucial to this
168 work, we also found a main effect of training day on the mean responses ($x^2_{df=1}$= 6.72, p
169 = 0.009, Fig. 3c), but not on the variance of these component amplitudes ($x^2_{df=1}$= 2.76, p
170 = 0.1, Fig. 3d). No significant interaction effects of sensory evidence and training day
171 were observed ($x^2_{df=1}$= 0.03, p = 0.86 and $x^2_{df=1}$= 0.25, p = 0.61, means and variance
172 respectively). These results suggest that the improvements in discrimination
173 performance for the Late component over the course of training are primarily the result
174 of gain modulation (i.e. enhanced sensory readout leading to amplification of the
175 differential response) of the component amplitudes rather than a reduction in the trial-to-
176 trial variability in these amplitudes.
177
178 To establish a concrete link between our EEG component amplitudes and improvements
179 in behaviour we ran a separate logistic regression analysis whereby trial-by-trial changes
180 in the amplitudes (i.e. y's) of the Early and Late components over all training days were
181 used to predict participants' choices on individual trials (i.e. face choice probability,

coded as 1 (0) for face (car) choices, respectively). Using the resulting subject-specific regression coefficients we found that our Late component was both a reliable predictor of participants' choices (t-test, $t(13) = 11.52$, $p < 0.001$) and a significantly better predictor compared to the Early component (paired t-test, $t(13) = 2.949$, $p = 0.011$).

Though the novelty of our work rests primarily with the EEG results, we also test the view that the observed perceptual improvements in behaviour might involve a RL-like mechanism similar to that proposed for reward-based learning [14,17,26,32]. To this end, we modelled our participants' choices using a RL model (see Methods). In brief, the model makes choices based on a decision variable (DV), with positive values indicating a higher likelihood of a face choice and negative values indicating a higher likelihood of a car choice. The DV reflects the representational strength of the presented stimulus on a given trial and corresponds to the stimulus sensory evidence scaled by the absolute difference between its signal weight and a noise weight for the antagonistic stimulus. Whilst the role of the former is to enhance the sensory read-out of the presented stimulus, the latter captures the extent to which the antagonistic stimulus interferes with the processing of the available sensory evidence.

In the RL framework employed here these weights are updated by means of a prediction error signal, which quantifies the discrepancy between the expected and actual value of the decision outcome on each trial. To account for the possibility that signal and noise weights may be differentially updated the prediction error signal is scaled by separate learning rates in each of the two weight updates. The mechanism of this update is such that on a given trial a correct choice will always lead to an increase of the chosen stimulus signal weight and to a decrease of the unchosen stimulus noise weight, yielding enhanced signal to noise ratio for the correctly chosen stimulus. Crucially this update is

also scaled by the chosen stimulus representation, which exerts a further consolidating effect on perceptual learning (see Methods).

We fit the model to individual participant data and found a highly significant correspondence between the model's accuracy predictions and actual behaviour ($r = 0.882$, $p < 0.001$ – Fig. 4a). We also compared the model with two competing alternatives (i.e. a model with signal and noise weights updated with only one learning rate and a model with only a single perceptual weight) using Bayesian Model Selection (BMS) that accounts for inter-subject variability by treating each model as a random effect. We found that our model provided a better fit to the observed choice behaviour (see Methods and Figure 4a). Consistent with an enhanced readout of sensory evidence we observed a subject-wise gradual build-up in the trial-by-trial estimates of the signal weights mirrored by a gradual decrease in the noise weight estimates (e.g. Fig. 4b and 4c respectively). Between-day comparisons (1 vs. 2 and 2 vs. 3) of subject-wise mean DVs (Fig. 4d; paired t-test: $t_{1vs2}\,(13) = -6.77$ $p < 0.001$; $t_{2vs3}\,(13) = -2.36$ $p = 0.02$) and aggregate perceptual weights (Fig. 4e; signal weights: paired t-test: $t_{1vs2}\,(13) = -6.74$ $p < 0.001$; $t_{2vs3}\,(13) = -2.36$ $p = 0.02$; noise weights: paired t-test: $t_{1vs2}\,(13) = 6.74$ $p < 0.001$; $t_{2vs3}\,(13) = 2.35$ $p = 0.02$) revealed a significant effect of learning as observed in behaviour.

To offer neurobiological validity to the model we performed two additional analyses. Firstly, we correlated the single-trial DVs estimated by the model with our EEG component amplitudes. We predicted that if the brain computes a version of our model-based DVs to drive choices then one should observe a systematic amplification of the DV with training and a significant correlation with our Late EEG component shown to index decision evidence. To this end we ran another regression analysis whereby the

single-trial amplitudes of our Early and Late components were used to predict the model's DVs. We found that our Late component was both a reliable predictor of the model's DVs (Fig. 4f; t-test, $t(13) = 21.81$, $p < 0.001$) and a significantly better predictor than the Early component (Fig. 4f; paired t-test, $t(13) = 3.06$, $p = 0.009$).

Secondly, we separated our trials into four bins (quartiles) based on the model-predicted magnitudes of the prediction error (PE) signal, which is thought to guide learning. We then ran a single-trial discriminant analysis on feedback-locked EEG data between the very low and very high PE trial groups (i.e. we kept the middle two quartiles as "test" data – see below). This analysis revealed a centroparietal EEG component peaking on average at 354ms post-feedback (Fig. 5a). The timing and topography of this component are consistent with previous work on feedback-related processing in the human brain using a probabilistic reinforcement learning task [33,34].

To formally test whether this EEG component was parametrically modulated by the magnitude of the PE signal, we computed discriminator amplitudes (y) for trials with intermediate magnitude levels (i.e. those left out from the original discrimination analysis). Specifically, we applied the spatial filter of the window that resulted in the highest discrimination performance for the extreme PE magnitude levels to the EEG data with intermediate values. We expected these "unseen" trials would show a parametric response profile such that the resulting mean component amplitude at the time of peak discrimination would proceed from very low < low < high < very high PE magnitude. Using this approach, we demonstrated that the mean discriminator output for each quartile increased as a function of the model's PE magnitude (all pair-wise t-test comparisons across adjacent trial groups: P values < 0.001; Fig. 5b), thereby establishing a concrete link between the model's PE estimates and our feedback-related

EEG component. Taken together, these findings provide further evidence that perceptual learning enhances decision-related evidence, likely via a RL-like mechanism.

Discussion

In this work, we offer the evidence from time-resolved electrophysiological signals in humans linking perceptual learning with post-sensory processing during a perceptual categorization task. Specifically, we showed that improvements in behavioural performance were accompanied primarily by late enhancements in decision-related evidence. In particular, we demonstrated that single-trial amplitudes of a late EEG component indexing decision evidence[19,20,23,35] were amplified in the course of learning, such that these representations became more robust to noise (rather than a reduction in noise as such). In contrast a temporally earlier component encoding sensory (stimulus) evidence – even in the absence of a face/car decision task[19] – was not affected by training. These findings suggest that it is the strengthening of the connections between early sensory encoding and downstream decision-related processing that are driving perceptual learning in our task.

Crucially, we also showed that the onset of the late component (which on average coincides with the onset of decision evidence accumulation [36-38]) systematically moves earlier in time with training. This finding is particularly interesting since we have previously observed comparable temporal shifts in this component while manipulating task/stimulus difficulty [19,20,23]. We view this as additional evidence that our learning effects on the late component lead to changes in perceptual sensitivity. More specifically, the earlier the onset time of the late component the stronger the behavioural improvements, consistent with a decrease in perceived task difficulty. These temporal

changes are also in line with a faster and more efficient accumulation of evidence as often predicted by sequential sampling models of decision making [29-31] (e.g. increases in the drift-rate and decrease in nondecision time variability).

Consistent with previous accounts [14,17] we also showed that these learning-induced behavioural improvements could be reliably explained in terms of a RL mechanism (see, e.g., [39]). More specifically, we showed that a model that uses a prediction error signal [24,25,40,41] to continuously adjust the stimulus specific perceptual weights on the sensory evidence [26] led to amplification of the relevant stimulus representations in the course of training (i.e. making them more robust to noise). We further demonstrated that trial-by-trial changes in our Late EEG component shown to index decision evidence reliably tracked the amplification of sensory information predicted by the model. These results imply that perceptual learning involves an enhanced readout of sensory information during decision making likely via a RL-like process, endorsing the view of a domain-general learning mechanism [24]. It is worth noting that whilst it is true that our task did not involve any explicit reward as a reinforcer, we view the implicit rewarding nature associated with correct responses as a “teaching signal” for strengthening the neural representation of sensory contingencies [26].

Research on perceptual learning has recently focused on the extent to which perceptual learning is due to improvements in sensory abilities that are (informationally and temporally) earlier than the decision process itself or due to improvements in post-sensory and decision-related processing. Consistent with the former account, several psychophysics studies have demonstrated that perceptual learning is often highly specific to the location and other properties of the stimuli [3-9], implying specificity to the trained retinal location [42,43]. Similarly human fMRI studies offered evidence of activity

enhancements in retinotopic areas corresponding to the trained visual fields [44] and increased responses along the whole hierarchy of early visual areas that correlated with improvements in behavioural performance following training over the course of several weeks [45,46]. These results are further corroborated by EEG recordings in humans showing post-training improvements in early visually-evoked components over occipital electrode sites [47-49] and electrophysiological recordings in NHPs linking behavioural performance with improvements in perceptual sensitivity in primary sensory areas [50-52].

In contrast, other psychophysical studies proposed that perceptual learning can also arise from changes in how sensory signals are read out or interpreted by decision-making mechanisms [32,53,54] rather than from changes in primary sensory areas as such. Neural evidence in support of this interpretation comes from NHP electrophysiology studies [13,14], demonstrating that perceptual learning on a motion discrimination task affects downstream decision accumulator areas, rather than regions encoding the sensory evidence (i.e. motion direction). Specifically, accumulator neurons improved responsiveness to the decision evidence in the course of learning (as reflected in steeper evidence accumulation slopes), with these improvements being proportional to the animals' performance on the task. Correspondingly, recent fMRI studies in humans started to explore the effect of learning on the activity and connectivity patterns of higher-level ventral temporal [55,56] and decision-related regions [15-18].

These seemingly discrepant accounts of the temporal locus of perceptual learning may be reconciled by considering differences in the experimental demands of the task at hand. For example, a recent theoretical account proposed a unified two-stage model of perceptual learning [57-59]. According to this model, there are two distinct types of plasticity underlying perceptual learning: feature-based plasticity and task-based plasticity. On the

one hand, feature based plasticity affects early sensory processing stages and occurs with mere exposure to stimuli, regardless of whether the stimuli are relevant to the task or not. Task-based plasticity, on the other hand, can be thought of as a higher-level processing stage arising from direct and active involvement in a behavioural task. In this formulation, the relative contribution of the two plasticity types to the overall enhancement in performance hinges largely on the training procedures, the stimuli and the intricacies of the task used in learning [60].

More specifically, a distinction could be drawn between tasks that involve learning of relatively primitive stimulus features such as orientation, spatial frequency or contrast and those employing more complex stimuli such as objects and faces [59]. Although learning of highly primitive features could occur locally at the level of early sensory processing, more complex stimuli (made up of a combination of primitive features) might require active involvement of downstream higher-level sensory or decision-related areas. In our design, for instance, complex object categories are used and phase discrimination, which is shown to involve processes beyond early visual cortex [61], is required to perform the task reliably. As such, our findings appear to rely heavily on the enhancement of the relevant stimulus representations during post-sensory, rather than early sensory processing.

In summary, our study provides critical insights into the neurobiology of perceptual learning and offers strong support to the notion that neuronal plasticity can occur at multiple time-scales and locations, depending on task demands and context. As such our findings can help revise existing theories of perceptual learning focusing only on early sensory processing and provide the foundation upon which future studies continue to interrogate the neural systems underlying perceptual decision making.

Methods

Participants

Fourteen subjects (7 female and 7 male, age range 23-28 years) participated in this study. All were right handed, had normal or corrected-to-normal vision and reported no history of neurological problems. The study was approved by the College of Science and Engineering Ethics Committee at the University of Glasgow (CSE01353) and informed consent was obtained from all participants.

Stimuli

We used a set of 18 face and 18 car images (image size 512 x 512 pixels, 8-bits/pixel), adapted from our previous experiments [19,20]. Face images were selected form the Face Database of the Max Planck Institute of Biological Cybernetics [62] and car images were sources from the internet. Both image types contained equal numbers of frontal and side views (up to ±45). All images were equated for spatial frequency, luminance and contrast and they all had identical magnitude spectra (average magnitude spectrum of all images in the database). We manipulated the phase spectra of the images using the weighted mean phase [63] technique to change the amount of sensory evidence in the stimuli as characterized by their % phase coherence. We selected two levels of sensory evidence for this study (32.5% and 37.5 % phase coherence) that are known to yield performance spanning psychophysical threshold, based on our previous studies [19,20]. A Dell Precision Workstation (Intel Core 2 Quad) running Windows 7 (64 bit) with an ATI FirePro 2270 graphics card and PsychoPy 1.8 presentation software [64] controlled the stimulus display. Images were presented on a Dell 2001FP TFT monitor (resolution, 1600x1200 pixels; refresh rate, 60 Hz). Subjects were positioned 75cm from the monitor and each image subtended approximately 6 x 6 degrees of visual angle.

Behavioural task

Subjects performed a simple image categorization task whereby they had to classify an image either as a face or car. The stimulus was presented for 50 ms and subjects were asked to make a response as soon as they had formed a decision, with a response deadline set at 1.25 s. Subjects indicated their decision with a button press on a response device (Cedrus RB-740) using their right index and middle fingers for a face and a car response, respectively. Subjects received visual feedback following each response that lasted for 500 ms. A tick and a cross were presented for a correct and an incorrect response, respectively (subtended 0.7 x 0.7 degrees of visual angle). A cross was also shown when subjects failed to make a response within the pre-allocated duration of 1.25 s following the stimulus. Feedback was followed by an inter-trial interval that varied randomly in the range between 1 – 1.5 s. There were a total of 288 trials (divided equally between the two image categories and the two levels of sensory evidence), presented in 4 blocks of 72 trials with a 60 s rest period between each block. The entire experiment lasted approximately 20 minutes. Each subject performed this task on three consecutive days, with the experiment taking place at the same time on each day. On the first day, subjects performed a short practice session of the face/car categorization task with high % phase coherence stimuli (50%) to familiarize themselves with the structure and pace of the task.

EEG data acquisition

EEG was collected inside an electrostatically shielded booth using a 64-channel EEG amplifier system (BrainAmps MR-Plus, Brain Products, Germany) and recorded using Brain Vision Recorder (BVR; Version 1.10, Brain Products, Germany) with a 1000 Hz sampling rate and an analogue bandpass filter of 0.016-250 Hz. The EEG cap consisted

of a 64 Ag/AgCl actiCAP electrodes (Brain Products, Germany) positioned according to the international 10–20 system of electrode positioning. The ground electrode was embedded in the EEG cap and placed along the midline between electrode Pz and Oz. The reference electrode was placed on the left mastoid. All input impedances were kept below 10 kΩ. For each participant, an effort was made to position the EEG cap in a consistent manner across the three training days, by keeping the distance between electrodes and certain anatomical landmarks (i.e. outer canthi, inion, nasion) constant. Experimental event codes and button responses were also synchronized with the EEG data and collected using the BVR software.

EEG pre-processing

We performed basic pre-processing of the EEG signals offline using Matlab (Mathworks, Natick, MA). Specifically, we applied a 0.5 Hz high-pass filter to remove DC drifts, and 100 Hz low pass filter to remove high frequency artefacts not associated with neurophysiological processes. These filters were applied together, non-causally to avoid distortions caused by phase delays (using MATLAB “filtfilt”). The EEG data was additionally re-referenced to the average of all channels.

Eye-movement artefact removal

Prior to the main experiment, we asked our participants to complete an eye movement calibration task during which they were instructed to blink repeatedly upon the appearance of a fixation cross in the centre of the screen and then to make several horizontal and vertical saccades according to the position of the fixation cross. The fixation cross subtended 0.4 x 0.4 degrees of visual angle. Horizontal saccades subtended 15 degrees and vertical saccades subtended 10 degrees. This exercise enabled us to determine linear EEG sensor weightings corresponding to eye blinks and

saccades (using principal component analysis) such that these components were projected onto the broadband data from the main task and subtracted out [27].

Single-trial discriminant analysis

To discriminate between face and car trials we applied a linear multivariate classifier to stimulus-locked EEG data, using the sliding window approach we used in previous work (e.g., [20,65]). Specifically, we identified a projection of the multichannel EEG signal, $\boldsymbol{x}_i(t)$, where *i* = [1…*T*] and *T* is the total number of trials, within a short time window that maximally discriminated between the two stimulus categories. All time windows had a width of N = 50 ms and the window centre τ was shifted from -100 to 1000 ms relative to stimulus onset, in 10 ms increments. More specifically, we used logistic regression [27] to learn a 64-channel spatial weighting, $\boldsymbol{w}(\tau)$, that achieved maximal discrimination at each time window, arriving at the one-dimensional projection $y_i(\tau)$, for each trial *i* and a given window τ:

$$y_i(\tau) = \frac{1}{N}\sum_{t=\tau-N/2}^{t=\tau+N/2} \boldsymbol{w}(\tau)^{\perp}\boldsymbol{x}_i(t) \quad (1)$$

where ⊥ is used to indicate the transpose operator. Note that our classifier is designed to return activity from processes that help maximize the difference across the two conditions of interest while minimizing the effect of processes common to both conditions. In doing so the classifier tries to map positive and negative discriminant component amplitudes (i.e. $y_i(\tau)$) to face and car trials, respectively. In other words, large positive values indicate a higher likelihood of a face stimulus, large negative values a higher likelihood of a car stimulus and values near zero reflect more difficult stimuli (see Fig. 2a for an example). This procedure in effect scales the resulting discriminating

component amplitudes in a manner that is directly comparable across the three training days. The same discrimination procedure was also applied on feedback-locked data to discriminate between trials with low versus high prediction error magnitudes (as estimated by our model – see below).

We quantified the performance of the discriminator at each time window using the area under a receiver operating characteristic (ROC) curve, referred to as an *Az*-value, using a leave-one-out trial procedure [66]. Furthermore we used a bootstrapping technique to assess the significance of the discriminator by performing the leave-one-out test after randomizing the trial labels. We repeated this randomization procedure 1000 times to produce a probability distribution for *Az*, and estimated the *Az* value leading to a significance level of $p < 0.01$.

Given the linearity of our model we also computed scalp topographies of the discriminating components resulting from Eq. 1 by estimating a forward model as:

$$a(\tau) = \frac{\boldsymbol{x}(\tau)\boldsymbol{y}(\tau)}{\boldsymbol{y}(\tau)^{\perp}\boldsymbol{y}(\tau)} \quad (2)$$

where $y_i(\tau)$ is now organized as a vector $\boldsymbol{y}(\tau)$, where each row is from trial *i*, and $\boldsymbol{x}_i(t)$ is organized as a matrix, $\boldsymbol{x}(\tau)$, where rows are channels and columns are trials, all for time window τ. These forward models can be viewed as scalp plots and interpreted as the coupling between the discriminating components and the observed EEG [27].

Single-trial regression analyses

To analyse the behavioural and neural data resulting from our EEG discrimination

analysis we use a mixed-effects general linear modelling (GLM) approach. These GLM models are similar to repeated-measures ANOVA models but they offer a better account for inter-subject response variability (by incorporating subjects as a random effect) and allow the mixing of both continuous and categorical variables [67,68]. Details of the dependent and predictor variables used for each regression analysis are given in the main text. The significance of a predictor variable or set of variables is tested using a log-likelihood ratio test, whereby the log-likelihood of the model with all predictors is compared to the log-likelihood of the model without the predictors being tested. The difference in the log-likelihood of two models is distributed according to a χ^2 distribution whose degrees of freedom equal the difference in the number of parameters in the two models. We fit these mixed-effects models using the lme4 package (http://cran.rproject.org/web/packages/lme4/index.html) using R (http://www.r-project.org). We note that repeating these analyses using a conventional ANOVA approach yielded virtually identical results, further highlighting the robustness of our effects.

To demonstrate that our Late EEG component was a better predictor of behaviour compared to the Early one, we ran a separate logistic regression analysis. Specifically, for each participant the trial-by-trial discriminant amplitudes (y's) for the two components (over all training days) were used as separate regressors to predict each participant's face choice probabilities ($P(f)$) on individual trials (i.e. $P(f)$ = 1 [0] for face [car] choices) as:

$$P(f) = 1/(1 + e^{-[\beta_0+\beta_1 y_{Early}+\beta_2 y_{Late}]}) \quad (3)$$

Then, to establish a more reliable trial-by-trial association between Late brain activity and choice behaviour, we tested: 1) whether the Late regression coefficients across subjects (β_2's) come from a distribution with mean greater than zero (using an one sample *t-test*) and 2) whether the Late regression coefficients across subjects (β_2's) come from a distribution with mean greater than those of the Early one (β_1's) (using a paired *t-test*). For all analyses, we provide exact p values where possible, but values below 10^{-3} are abbreviated as such (i.e. $p < 0.001$).

Reinforcement learning model

We used a variant of the Rescorla-Wagner reinforcement learning model to account for perceptual improvements in the course of learning [14,17,26]. In this model perceptual decisions are driven by a decision variable (DV) denoting the subject's hidden representations of sensory contingencies (i.e. association between sensory evidence and stimulus category). The strength of such representations is modulated via dynamic updates of category specific perceptual weights based on feedback information, thereby accounting for potential differences in learning trajectories between the stimulus categories. Indeed, compared to previous work that used a single stimulus-invariant perceptual weight [14,17] the introduction of category-specific perceptual weights is designed to capture subject-wise choice biases in that, subjects might have a choice bias towards cars or faces and likewise might display an increasing ability to recognize cars or faces throughout the task.

Moreover, our perceptual weights comprise signal and noise weights. Whilst the former is designated to enhance stimulus representation in the course of learning, the latter accounts for the interference exerted by the antagonistic stimulus against the acquisition of the correct sensory contingencies. Thus, in our model perceptual learning is expected

to occur via gradually increasing signal weights as well as gradually decreasing noise weights. Compared to previous RL-like perceptual models[14,17], this better captures instances whereby improved task performance depends both on greater ability to recognize a given stimulus as well as greater ability to rule out the antagonistic stimulus. In other words on a face trial subjects might correctly choose face partly because they are able to identify face-like features and partly because they are able to recognize that there are no car-like features.

More specifically, on each trial i decision activities specific to each stimulus category ($A_{stim}, stim \in \{face, car\}$) were estimated as the stimulus specific sensory evidence (E_{stim}) scaled by the absolute difference between the stimulus specific signal weight (v_{stim}) and the noise weight of the antagonistic stimulus ($n_{\backslash stim}$):

$$A^i_{face} = E^i_{face}|v^i_{face} - n^i_{car}|$$
$$A^i_{car} = E^i_{car}|v^i_{car} - n^i_{face}| \quad (4)$$

As perceptual learning progresses, the estimates of signal and noise weights grow apart and so does their distance (i.e. absolute difference) on the real line. As a result, the read-out of sensory evidence is increasingly enhanced reflecting the improving ability to discriminate between perceptual stimuli in the course of training.

Whilst the magnitude of E_{stim} was defined according to the percentage of phase coherence in the stimulus (0.325 and 0.375 for low and high coherence trials respectively), its sign was related to stimulus category (positive for faces and negative for cars). This ensured decision activities to be a signed quantity, whose magnitude

tracked the time-varying strength of stimulus representation.

Trial-by-trial estimates of the DV were computed based on the decision activity of the presented stimulus:

$$DV^i = A^i_{stim} \quad (5)$$

Note that the DV too is a signed quantity with positive values indicating a higher likelihood of a face choice and negative values indicating a higher likelihood of a car choice and as such is directly comparable with the sign of our EEG discriminator component amplitudes y. Correspondingly, both the model's DVs and our component amplitudes are orthogonal to potentially confounding quantities such as task (stimulus) difficulty, decision confidence (or uncertainty) and expected value, all of which covary with the absolute value of the DV and y (i.e. both high positive and high negative DV and y values correspond to easier, more confident and thereby higher expected value choices).

Subject-wise DV trajectories were then mapped to choice propensities (i.e. probabilities) using a sigmoid function:

$$P^i_{face} = \sigma(\beta(DV^i)) \quad (6)$$

where $\sigma(z) = 1/(1 + e^{-z})$ is the sigmoid function and β the inverse of the temperature representing the degree of stochasticity in the decision function. Next, the expected value (EV) of the outcome on the same trial was computed based on the modulus

(absolute value) of the DV as: $EV^i = \sigma(\beta(|DV^i|))$. In other words, whilst high positive and high negative DV values (subjectively easier choices) increase the EV of the outcome (and therefore the expected probability of being rewarded), values near zero (subjectively difficult choices) reduce it.

Finally, on each trial, given reward feedback r (coded as 1 and 0 for reward and no reward respectively), perceptual weights were updated via a prediction error signal, $\delta^i = r^i - EV^i$, which quantified the degree of deviation between the actual and expected outcome, scaled by a learning rate parameter α and an associativity component ($E^i_{stim}A^i_{choice}$) whose role was to dynamically modulate the updating of perceptual weights depending on the strength of sensory evidence (E_{stim}) and strength of the chosen stimulus representation (A_{choice}):

$$v^{i+1}_{choice} = v^i_{choice} + \alpha\delta^i E^i_{choice}A^i_{choice}$$
$$n^{i+1}_{\backslash choice} = v^i_{\backslash choice} + \alpha\delta^i E^i_{\backslash choice}A^i_{choice} \quad (7)$$

where subscript $\backslash choice$ indicates the unchosen stimulus. Note that the signal weight of the unchosen stimulus and the noise weight of the chosen stimulus were not updated. The sign of the update was determined by the prediction error so that whilst correct choice trials resulted in an increase of signal weights and a decrease of noise weights, incorrect choice trials had an opposite effect on the updating of perceptual weights. For example on a face trial whilst a correct face choice would result in an increase of v_{face} and a reduction of n_{car} an incorrect car choice would yield an increase of n_{car} and a reduction of v_{face}. Moreover, the learning/unlearning of correct/incorrect sensory contingencies underpinned by this dynamic updating of perceptual weights was further

facilitated by the strength of the stimulus representation A_{choice}. In other words, the stronger the stimulus representation the greater the impact of the prediction error on perceptual learning (via the updating of perceptual weights) and vice versa.

We fitted two variants of this model, one with a single learning rate and one with two different learning rates for the signal and noise weights' updates, respectively. The latter model allowed for the possibility that signal and noise weights may be differentially updated therefore probing subject-specific biases in perceptual information processing. In other words, whilst some subjects might boost the signal to noise ratio for a given stimulus by primarily enhancing signal weights, others might achieve the same result by primarily reducing noise weights. Whilst the total number of free parameters in the first variant of the model was four $(v^1, n^1, \beta, \alpha)$ the number of free parameters in the second variant was five $(v^1, n^1, \beta, \alpha_{face}, \alpha_{car})$ where v^1, n^1 represent the initial perceptual weights' estimates (i.e. on the first trial) for face and car stimuli. In addition, we also fitted a simple perceptual RL model (as described in [14,17]), whereby the read-out of sensory evidence was scaled only by a signal weight and the trial-by-trial updating of this signal weight was driven by a prediction error computed as previously illustrated. The number of free parameters in this model was four $(v^1, bias, \beta, \alpha)$ where $bias$ represents the indecision point in the choice sigmoid function.

Model fitting procedure and model comparison

In order to prevent overfitting, for each subject i we found the *maximum a posteriori* estimate of the model free parameters:

$$\theta_i^{MAP} = argmax_\theta \, p(C_i|\theta_i)p(\theta_i|\xi) \quad (8)$$

where $p(C_i|\theta_i)$ is the cross-entropy loss function between empirical and predicted choices C_i given the model parameters θ_i and $p(\theta_i|\xi)$ is the prior distribution on the model parameters θ_i given the population-level hyperparameters ξ. Priors were defined as normal distributions $N(\mu, \sigma)$ where μ was sampled from a normal distribution with mean 0 and standard deviation 1 and σ was set to 1e2. To preserve the parameters' natural bounds, log (β) and logit (α) transforms of the parameters were implemented.

We subsequently performed formal Bayesian model comparison between the 3 models to determine the one that best fitted our behavioural data. This approach treats each model as a random-effect at the between subject-level and therefore is more robust to outliers than fixed-effect approaches [69]. Specifically, we first estimated the subject-wise Laplace approximated log evidence for each model. We subsequently computed the model-wise exceedance probability (i.e. how confident we are that a model is more likely than any other model tested) using SPM8's spm_BMS routine [70]. We found that the exceedance probability of the model with 2 learning rates $(\varphi = 0.88)$ was greater than those of the model with a single learning rate $(\varphi = 0.11)$ and with a single perceptual weight $(\varphi = 0.001)$ (see inset Fig. 4a).

To assess the model's goodness of fit we plotted the subject-wise empirical choice accuracy against the model's predicted accuracy for different days and stimulus phase coherence levels. Additionally, we tested whether subject-wise model's mean *DV*s and perceptual weights significantly increased over training as observed with behavioural performance.

Data availability

The data that support the findings of this study are available from the corresponding author upon request.

References

1 Gilbert, C. D., Sigman, M. & Crist, R. E. The neural basis of perceptual learning. *Neuron* **31**, 681-697 (2001).

2 Goldstone, R. L. Perceptual learning. *Annual review of psychology* **49**, 585-612, doi:10.1146/annurev.psych.49.1.585 (1998).

3 Ball, K. & Sekuler, R. Direction-specific improvement in motion discrimination. *Vision research* **27**, 953-965 (1987).

4 Crist, R. E., Kapadia, M. K., Westheimer, G. & Gilbert, C. D. Perceptual learning of spatial localization: Specificity for orientation, position, and context. *Journal of Neurophysiology* **78**, 2889-2894 (1997).

5 Fahle, M. & Edelman, S. Long-term learning in vernier acuity: Effects of stimulus orientation, range and of feedback. *Vision research* **33**, 397-412 (1993).

6 Fiorentini, A. & Berardi, N. Perceptual learning specific for orientation and spatial frequency. *Nature* (1980).

7 Karni, A. & Sagi, D. Where practice makes perfect in texture discrimination: evidence for primary visual cortex plasticity. *Proceedings of the National Academy of Sciences* **88**, 4966-4970 (1991).

8 Poggio, T., Fahle, M. & Edelman, S. Fast perceptual learning in visual hyperacuity. *Science* **256**, 1018-1021 (1992).

9 Sagi, D. & Tanne, D. Perceptual learning: learning to see. *Current opinion in neurobiology* **4**, 195-199 (1994).

10 Ahissar, M. & Hochstein, S. Task difficulty and the specificity of perceptual learning. *Nature* **387**, 401-406, doi:10.1038/387401a0 (1997).

11 Mollon, J. D. & Danilova, M. V. Three remarks on perceptual learning. *Spatial vision* **10**, 51-58 (1996).

12 Ahissar, M. & Hochstein, S. Attentional control of early perceptual learning. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 5718-5722 (1993).

13 Law, C.-T. & Gold, J. I. Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area. *Nature neuroscience* **11**, 505-513 (2008).

14 Law, C.-T. & Gold, J. I. Reinforcement learning can account for associative and perceptual learning on a visual-decision task. *Nature neuroscience* **12**, 655-663 (2009).

15 Baldassarre, A. *et al.* Individual variability in functional connectivity predicts performance of a perceptual task. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 3516-3521, doi:10.1073/pnas.1113148109 (2012).

16 Chen, N. H. *et al.* Sharpened cortical tuning and enhanced cortico-cortical communication contribute to the long-term neural mechanisms of visual motion perceptual learning. *NeuroImage* **115**, 17-29, doi:10.1016/j.neuroimage.2015.04.041 (2015).

17 Kahnt, T., Grueschow, M., Speck, O. & Haynes, J.-D. Perceptual learning and decision-making in human medial frontal cortex. *Neuron* **70**, 549-559 (2011).

18 Lewis, C. M., Baldassarre, A., Committeri, G., Romani, G. L. & Corbetta, M. Learning sculpts the spontaneous activity of the resting human brain. *Proceedings of the National Academy of*

Sciences of the United States of America **106**, 17558-17563, doi:10.1073/pnas.0902455106 (2009).

19 Philiastides, M. G., Ratcliff, R. & Sajda, P. Neural representation of task difficulty and decision making during perceptual categorization: a timing diagram. *The Journal of Neuroscience* **26**, 8965-8975 (2006).

20 Philiastides, M. G. & Sajda, P. Temporal characterization of the neural correlates of perceptual decision making in the human brain. *Cerebral cortex* **16**, 509-518 (2006).

21 Philiastides, M. G. & Sajda, P. Causal influences in the human brain during face discrimination: a short-window directed transfer function approach. *IEEE transactions on bio-medical engineering* **53**, 2602-2605, doi:10.1109/TBME.2006.885122 (2006).

22 Philiastides, M. G. & Sajda, P. EEG-informed fMRI reveals spatiotemporal characteristics of perceptual decision making. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **27**, 13082-13091, doi:10.1523/JNEUROSCI.3540-07.2007 (2007).

23 Ratcliff, R., Philiastides, M. G. & Sajda, P. Quality of evidence for perceptual decision making is indexed by trial-to-trial variability of the EEG. *Proceedings of the National Academy of Sciences* **106**, 6539-6544 (2009).

24 Rushworth, M. F., Mars, R. B. & Summerfield, C. General mechanisms for making decisions? *Current opinion in neurobiology* **19**, 75-83 (2009).

25 Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science* **275**, 1593-1599 (1997).

26 Guggenmos, M., Wilbertz, G., Hebart, M. N. & Sterzer, P. Mesolimbic confidence signals guide perceptual learning in the absence of external feedback. *eLife* **5**, doi:10.7554/eLife.13388 (2016).

27 Parra, L. C., Spence, C. D., Gerson, A. D. & Sajda, P. Recipes for the linear analysis of EEG. *NeuroImage* **28**, 326-341, doi:10.1016/j.neuroimage.2005.05.032 (2005).

28 Sajda, P., Philiastides, M. G. & Parra, L. C. Single-trial analysis of neuroimaging data: inferring neural networks underlying perceptual decision-making in the human brain. *IEEE reviews in biomedical engineering* **2**, 97-109, doi:10.1109/RBME.2009.2034535 (2009).

29 Ratcliff, R. & Smith, P. L. Perceptual discrimination in static and dynamic noise: the temporal relation between perceptual encoding and decision making. *Journal of experimental psychology. General* **139**, 70-94, doi:10.1037/a0018128 (2010).

30 Ratcliff, R., Smith, P. L. & McKoon, G. Modeling Regularities in Response Time and Accuracy Data with the Diffusion Model. *Current directions in psychological science* **24**, 458-470, doi:10.1177/0963721415596228 (2015).

31 Smith, P. L. & Ratcliff, R. Psychology and neurobiology of simple decisions. *Trends in neurosciences* **27**, 161-168, doi:10.1016/j.tins.2004.01.006 (2004).

32 Petrov, A. A., Dosher, B. A. & Lu, Z.-L. The dynamics of perceptual learning: an incremental reweighting model. *Psychological review* **112**, 715 (2005).

33 Fouragnan, E., Retzler, C., Mullinger, K. & Philiastides, M. G. Two spatiotemporally distinct value systems shape reward-based

learning in the human brain. *Nature communications* **6**, 8107, doi:10.1038/ncomms9107 (2015).
34 Philiastides, M. G., Biele, G., Vavatzanidis, N., Kazzer, P. & Heekeren, H. R. Temporal dynamics of prediction error processing during reward-based decision making. *NeuroImage* **53**, 221-232, doi:10.1016/j.neuroimage.2010.05.052 (2010).
35 Lou, B., Li, Y., Philiastides, M. G. & Sajda, P. Prestimulus alpha power predicts fidelity of sensory encoding in perceptual decision making. *NeuroImage* **87**, 242-251, doi:10.1016/j.neuroimage.2013.10.041 (2014).
36 Kelly, S. P. & O'Connell, R. G. Internal and external influences on the rate of sensory evidence accumulation in the human brain. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **33**, 19434-19441, doi:10.1523/JNEUROSCI.3355-13.2013 (2013).
37 O'Connell, R. G., Dockree, P. M. & Kelly, S. P. A supramodal accumulation-to-bound signal that determines perceptual decisions in humans. *Nature neuroscience* **15**, 1729-1735, doi:10.1038/nn.3248 (2012).
38 Philiastides, M. G., Heekeren, H. R. & Sajda, P. Human scalp potentials reflect a mixture of decision-related signals during perceptual choices. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **34**, 16877-16889, doi:10.1523/JNEUROSCI.3012-14.2014 (2014).
39 Sutton, R. S. & Barto, A. G. *Reinforcement learning: An introduction*. (MIT press, 1998).
40 O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H. & Dolan, R. J. Temporal difference models and reward-related learning in the human brain. *Neuron* **38**, 329-337 (2003).
41 Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J. & Frith, C. D. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* **442**, 1042-1045 (2006).
42 Fahle, M. Perceptual learning: A case for early selection. *Journal of vision* **4**, 4-4 (2004).
43 Fahle, M. Perceptual learning: specificity versus generalization. *Current opinion in neurobiology* **15**, 154-160 (2005).
44 Schwartz, S., Maquet, P. & Frith, C. Neural correlates of perceptual learning: a functional MRI study of visual texture discrimination. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 17137-17142, doi:10.1073/pnas.242414599 (2002).
45 Furmanski, C. S., Schluppeck, D. & Engel, S. A. Learning strengthens the response of primary visual cortex to simple patterns. *Current Biology* **14**, 573-578 (2004).
46 Jehee, J. F., Ling, S., Swisher, J. D., van Bergen, R. S. & Tong, F. Perceptual learning selectively refines orientation representations in early visual cortex. *The Journal of neuroscience* **32**, 16747-16753 (2012).
47 Bao, M., Yang, L., Rios, C., He, B. & Engel, S. A. Perceptual learning increases the strength of the earliest signals in visual cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **30**, 15080-15084, doi:10.1523/JNEUROSCI.5703-09.2010 (2010).
48 Pourtois, G., Rauss, K. S., Vuilleumier, P. & Schwartz, S. Effects of perceptual learning on primary visual cortex activity

in humans. *Vision research* **48**, 55-62, doi:10.1016/j.visres.2007.10.027 (2008).
49 Censor, N., Bonneh, Y., Arieli, A. & Sagi, D. Early-vision brain responses which predict human visual segmentation and learning. *Journal of vision* **9**, 12 11-19, doi:10.1167/9.4.12 (2009).
50 Ghose, G. M., Yang, T. & Maunsell, J. H. Physiological correlates of perceptual learning in monkey V1 and V2. *Journal of neurophysiology* **87**, 1867-1888 (2002).
51 Schoups, A., Vogels, R., Qian, N. & Orban, G. Practising orientation identification improves orientation coding in V1 neurons. *Nature* **412**, 549-553, doi:10.1038/35087601 (2001).
52 Yan, Y. *et al.* Perceptual training continuously refines neuronal population codes in primary visual cortex. *Nature neuroscience* **17**, 1380-1387 (2014).
53 Dosher, B. A. & Lu, Z. L. Mechanisms of perceptual learning. *Vision research* **39**, 3197-3221, doi:Doi 10.1016/S0042-6989(99)00059-0 (1999).
54 Lu, Z.-L., Liu, J. & Dosher, B. A. Modeling mechanisms of perceptual learning with augmented Hebbian re-weighting. *Vision research* **50**, 375-390 (2010).
55 Kuai, S.-G., Levi, D. & Kourtzi, Z. Learning optimizes decision templates in the human visual cortex. *Current Biology* **23**, 1799-1804 (2013).
56 Li, S., Mayhew, S. D. & Kourtzi, Z. Learning shapes the representation of behavioral choice in the human brain. *Neuron* **62**, 441-452 (2009).
57 Shibata, K., Watanabe, T., Sasaki, Y. & Kawato, M. Perceptual learning incepted by decoded fMRI neurofeedback without stimulus presentation. *science* **334**, 1413-1415 (2011).
58 Shibata, K., Sagi, D. & Watanabe, T. Two-stage model in perceptual learning: toward a unified theory. *Annals of the New York Academy of Sciences* **1316**, 18-28, doi:10.1111/nyas.12419 (2014).
59 Watanabe, T. & Sasaki, Y. Perceptual learning: toward a comprehensive theory. *Annual review of psychology* **66**, 197-221, doi:10.1146/annurev-psych-010814-015214 (2015).
60 Li, W., Piech, V. & Gilbert, C. D. Perceptual learning and top-down influences in primary visual cortex. *Nature neuroscience* **7**, 651-657, doi:10.1038/nn1255 (2004).
61 Perna, A., Tosetti, M., Montanaro, D. & Morrone, M. C. BOLD response to spatial phase congruency in human brain. *Journal of vision* **8**, 15 11-15, doi:10.1167/8.10.15 (2008).
62 Troje, N. F. & Bülthoff, H. H. Face recognition under varying poses: The role of texture and shape. *Vision research* **36**, 1761-1771 (1996).
63 Dakin, S. C., Hess, R. F., Ledgeway, T. & Achtman, R. L. What causes non-monotonic tuning of fMRI response to noisy images? *Current biology : CB* **12**, R476-477; author reply R478 (2002).
64 Peirce, J. W. PsychoPy—psychophysics software in Python. *Journal of neuroscience methods* **162**, 8-13 (2007).
65 Gherman, S. & Philiastides, M. G. Neural representations of confidence emerge from the process of decision formation during perceptual choices. *NeuroImage* **106**, 134-143, doi:10.1016/j.neuroimage.2014.11.036 (2015).

66 Duda, R. O. & Hart, P. E. *Pattern classification and scene analysis*. Vol. 3 (Wiley New York, 1973).
67 Baayen, R. H., Davidson, D. J. & Bates, D. M. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* **59**, 390-412, doi:10.1016/j.jml.2007.12.005 (2008).
68 Gelman, A. & Hill, J. *Data analysis using regression and multilevel/hierarchical models*. (Cambridge University Press, 2006).
69 Stephan, K. E. *et al.* Ten simple rules for dynamic causal modeling. *NeuroImage* **49**, 3099-3109, doi:10.1016/j.neuroimage.2009.11.015 (2010).
70 Stephan, K. E., Weiskopf, N., Drysdale, P. M., Robinson, P. A. & Friston, K. J. Comparing hemodynamic models with DCM. *NeuroImage* **38**, 387-401, doi:10.1016/j.neuroimage.2007.07.040 (2007).

907 Correspondence and requests for materials should be send to Marios G. Philiastides
908 (marios.philiastides@glasow.ac.uk).
909
910 **Acknowledgements:** This work was supported by the Economic and Social Research
911 Council (ESRC; grant ES/L012995/1 to M.G.P.). The funder had no role in
912 conceptualization, research design, data collection, analysis, decision to publish or
913 preparation of the manuscript.
914
915 **Author contributions:** J.A.D. and M.G.P. designed the experiments. J.A.D. performed
916 the experiments. J.A.D., F.Q. and M.G.P. analyzed the data and wrote the paper.
917
918 **Competing interests:** The authors declare no competing interests.
919

Figure Legends

Figure 1. Experimental design and behaviour. **(a)** Schematic representation of the experimental paradigm. Subjects had to categorize a noisy image presented for 50ms as a face or a car and indicate their choice with a button press within 1250ms following the stimulus presentation. Feedback was then presented for 500ms (a tick or a cross for a correct and an incorrect response, respectively) followed by an inter-stimulus interval (ISI) that varied randomly between 1-1.5s. Subjects performed this task on three consecutive training days. Sample face and car images at the two levels of phase coherence used in the task (32.5% and 37.5%) are shown on the right. **(b)** Proportion of correct choices and **(c)** Mean reaction times (RT) as a function of the three training days (1: blue, 2: green, 3: red) and the two levels of phase coherence of the stimuli, averaged across subjects. Faint lines represent individual subject data. Error bars represent standard errors across subjects.

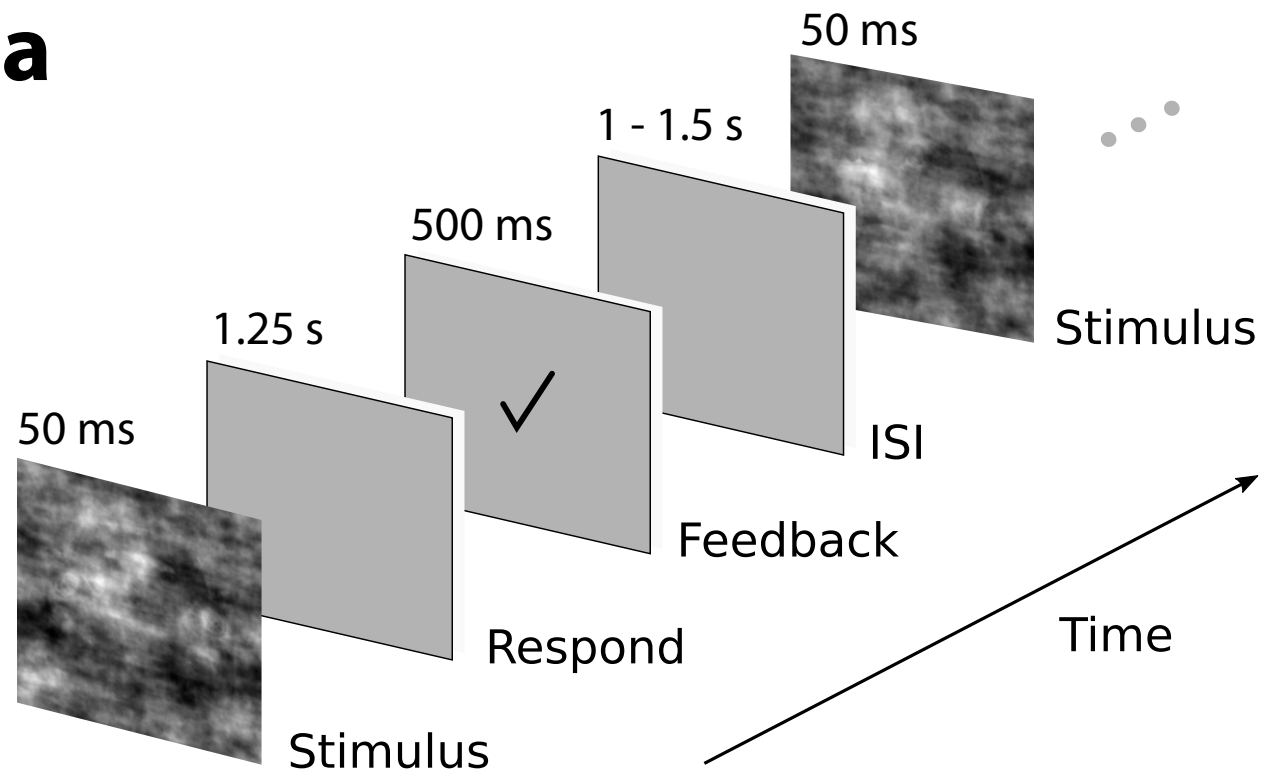
Figure 2. Post-sensory effects of perceptual learning. **(a)** Single-trial discriminator amplitudes (y) for the Early (dotted) and Late (solid) component windows for faces (black) and cars (grey) at 37.5% phase coherence from a representative subject on the third training day. The component amplitudes are shown as histograms on the right, with a cutoff (the thick black line) to separate trials into positive vs. negative amplitudes, indicating a higher likelihood of a face and a car trial, respectively. **(b)** Multivariate discriminator performance (A_z) during face-vs-car outcome discrimination of stimulus-locked EEG responses across the three training days (1: blue, 2: green, 3: red), averaged across subjects, showing the presence of the Early and Late components. The dotted line represents the average A_z value leading to a significance level of p = 0.01,

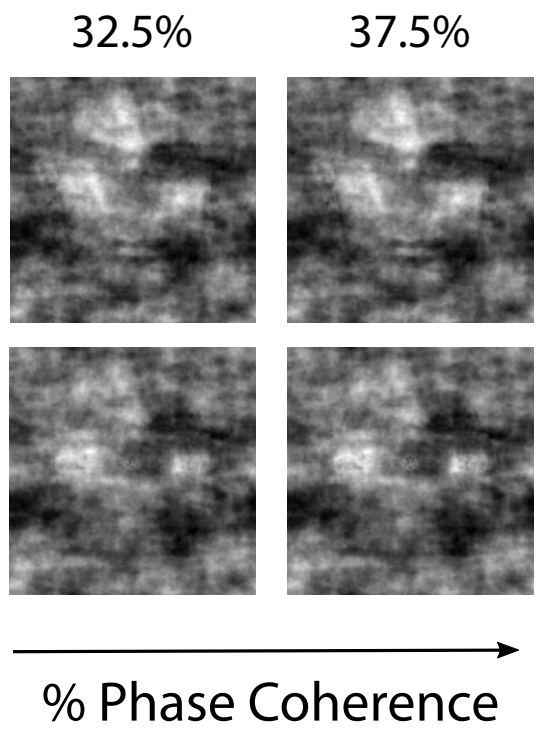
946 estimated using a bootstrap test. Faint lines represent individual subject data. **(c)**
947 Average discriminator performance and scalp topographies for the Early (magenta) and
948 Late (cyan) components across the three training days estimated at time of subject-
949 specific maximum discrimination. Faint lines represent individual subject data. Error bars
950 represent standard errors across subjects. **(d)** Average onset times for the Early
951 (magenta) and Late (cyan) components across the three training days. Faint lines
952 represent individual subject data. Error bars represent standard errors across subjects.
953
954 **Figure 3.** Enhanced readout of post-sensory decision evidence. **(a)** Schematic
955 illustration of possible effects on the distribution of single-trial discriminator amplitudes in
956 the course of learning. Top: increases in the distance between the mean response for
957 faces and cars. Bottom: reduction in the variance of the face and car responses. Both
958 examples lead to a smaller overlap (more separation) between the face and car
959 distributions. **(b)** Changes in the mean distance between the face and car distributions
960 for the Late component across the three training days (1: blue, 2: green, 3: red). **(c)**
961 Changes in the variance of the face and car distributions for the Late component across
962 the three training days. The faint lines in **(b)** and **(c)** represent individual subject data,
963 while the error bars represent standard errors across subjects.
964
965 **Figure 4.** Reinforcement learning model for perceptual choices. **(a)** Scatter plot showing
966 the correlation between the performance of individual subjects and models, over the
967 three training days and the two level of stimulus phase coherence (using the winning
968 model). Inset: exceedance probabilities of three competing models (see Methods for
969 details). **(b)** Individual trial estimates of the model's decision variable (DV) for a
970 representative subject over the course of the three training days, superimposed on the
971 amount of stimulus-defined sensory evidence (black trace). **(c)** Signal (positive) and

noise (negative) perceptual weights for faces (solid) and cars (dashed) over the three training days for the same subject shown in **(b)**. **(d)** Average magnitude of the model's DVs across subjects over the course of the three training days. Individual subject data are also shown as point estimates. **(e)** Average signal (positive) and noise (negative) perceptual weights for faces (solid) and cars (faint) over the three training days. Individual subject data are also shown as point estimates. **(f)** Average regression coefficients reflecting the trial-by-trial association between the model's DVs and the amplitudes of the Early and Late EEG components estimated over all training days. Individual subject data are also shown as point estimates.

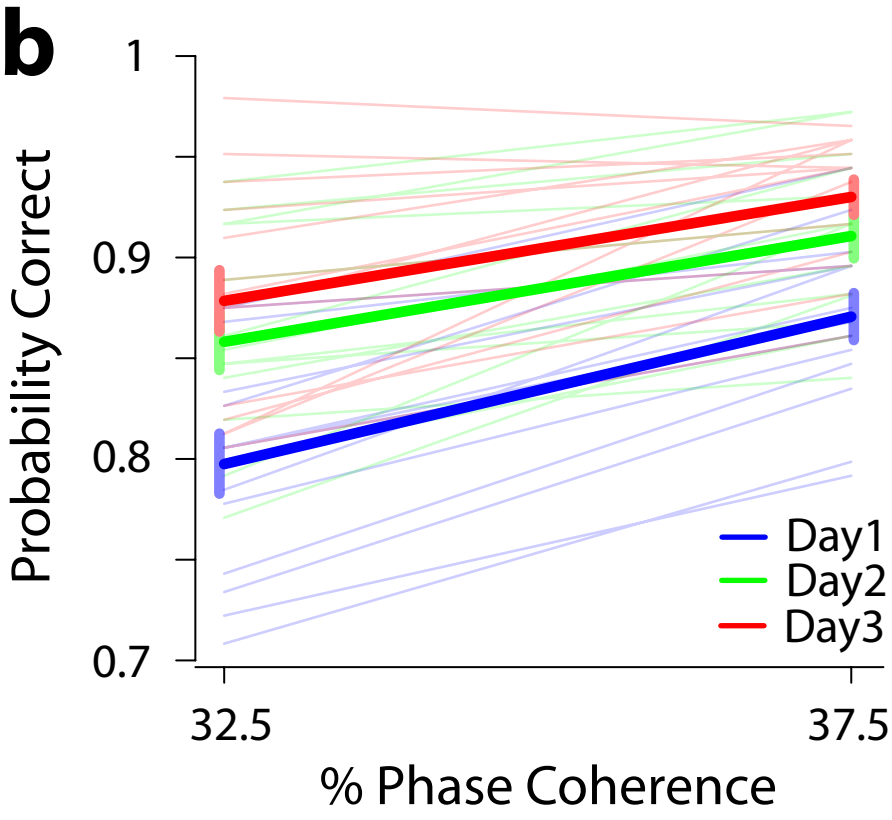
Figure 5. Electrophysiological correlates of prediction error (PE). **(a)** Multivariate discriminator performance (A_z) during very low versus very high PE magnitude trials on feedback-locked EEG responses averaged across subjects and days revealing a late PE component. Discriminator performance and component peak times were comparable across the three days. The dotted line represents the average A_z value leading to a significance level of p = 0.01, estimated using a bootstrap test. Faint lines represent individual subject data. Inset: average scalp topography associated with the PE component, estimated at time of subject-specific maximum discrimination. **(b)** Mean discriminator amplitude (y) for the PE component, binned in four quartiles based on model-based estimates of the magnitude of the PE, showing a clear parametric response along the four trial groups. Quartiles 1 and 4 were used to train the classifier, while quartiles 2 and 3 contain "unseen" data with intermediate PE magnitude levels. Individual subject data are also shown as point estimates.

a

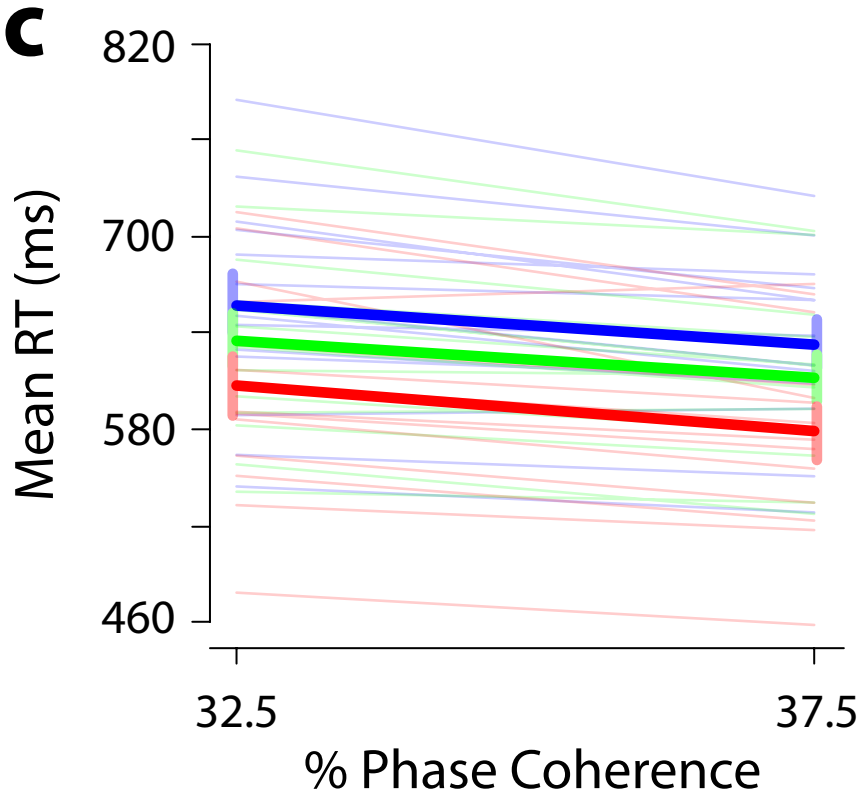

50 ms
1 - 1.5 s
500 ms
1.25 s
50 ms
Stimulus
ISI
Feedback
Respond
Stimulus
Time


32.5%
37.5%
% Phase Coherence

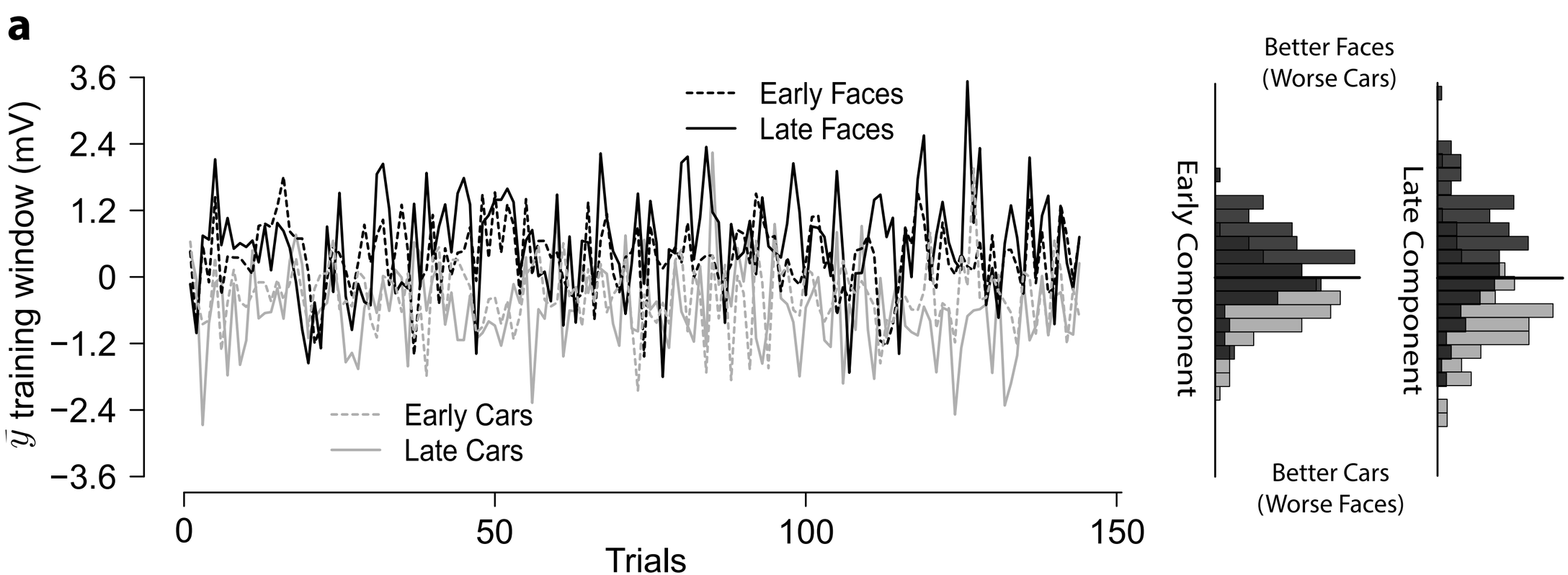
b


Probability Correct
1
0.9
0.8
0.7
Day1
Day2
Day3
32.5
37.5
% Phase Coherence

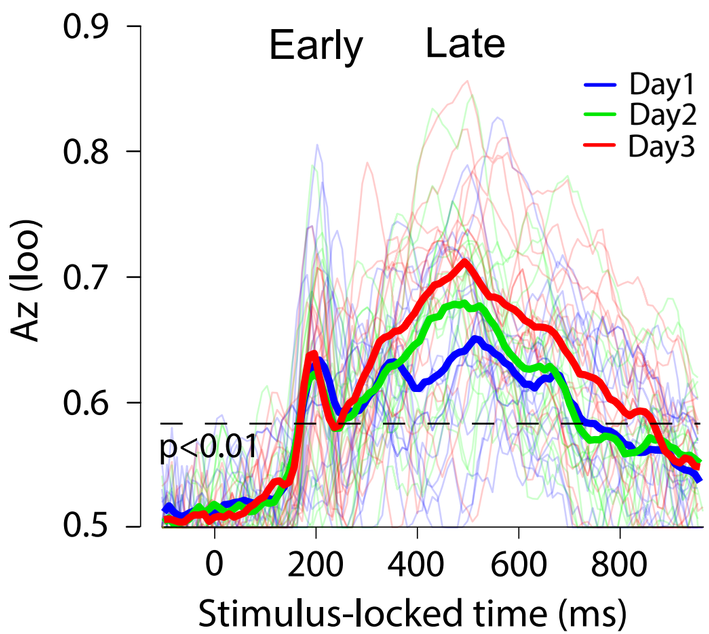
c


Mean RT (ms)
820
700
580
460
32.5
37.5
% Phase Coherence

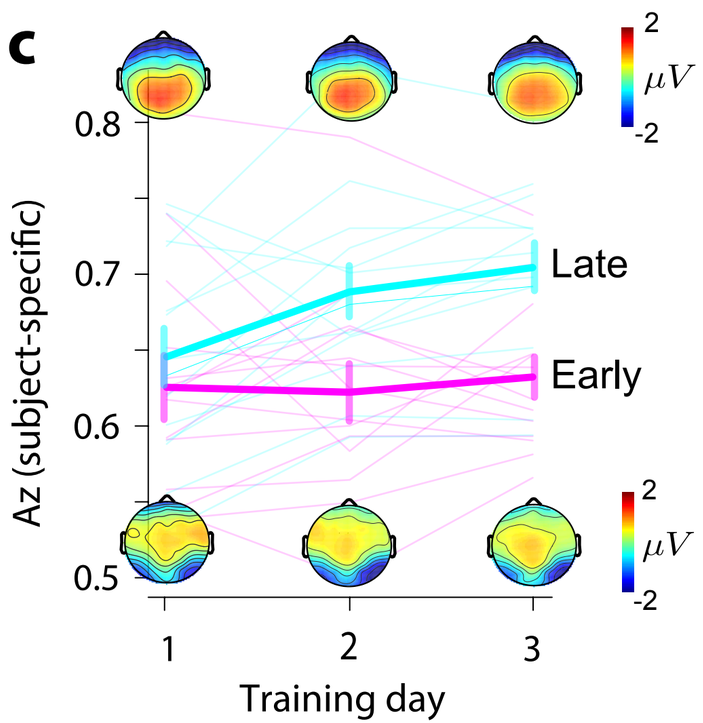
a


3.6
2.4
1.2
0
−1.2
−2.4
−3.6
$\bar{y}$ training window (mV)
Early Faces
Late Faces
Early Cars
Late Cars
0
50
100
150
Trials
Better Faces
(Worse Cars)
Early Component
Late Component
Better Cars
(Worse Faces)

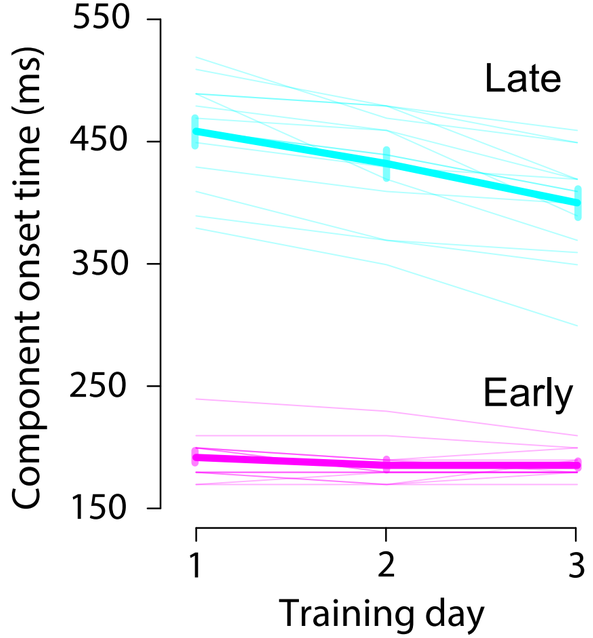
b


0.9
0.8
0.7
0.6
0.5
Az (loo)
Early
Late
Day1
Day2
Day3
p<0.01
0
200
400
600
800
Stimulus-locked time (ms)

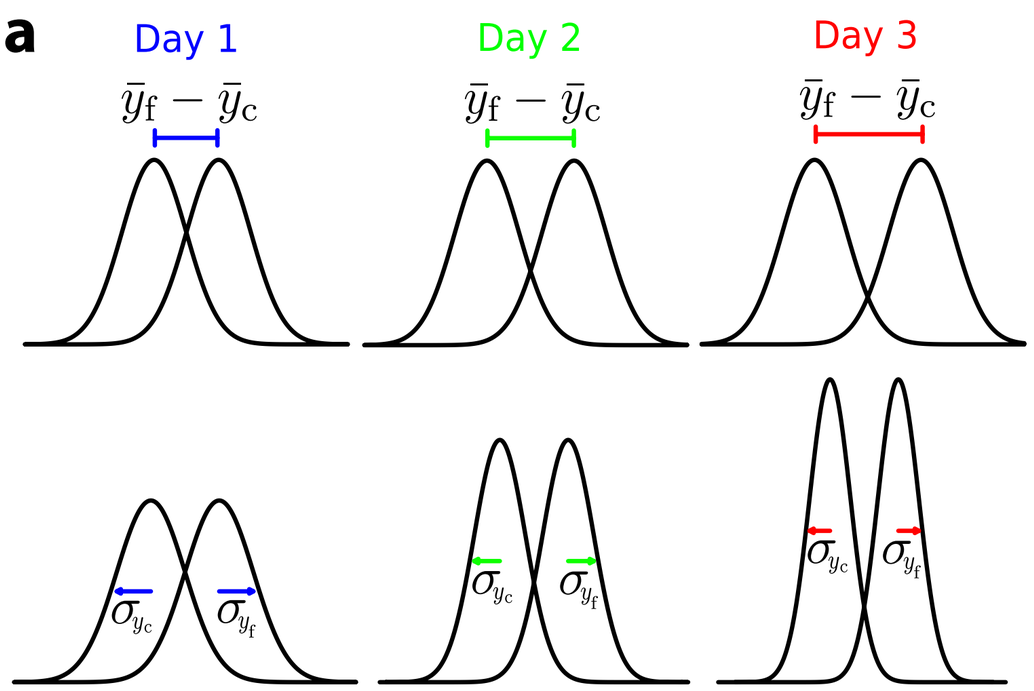
c


0.8
0.7
0.6
0.5
Az (subject-specific)
2
μV
-2
Late
Early
1
2
3
Training day

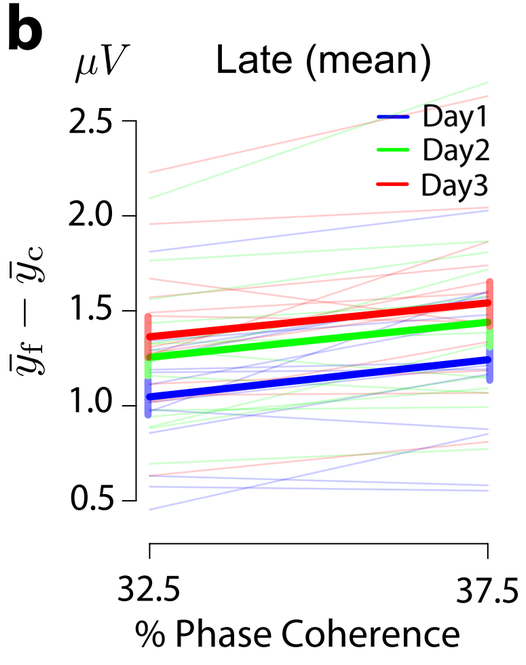
d


550
450
350
250
150
Component onset time (ms)
Late
Early
1
2
3
Training day

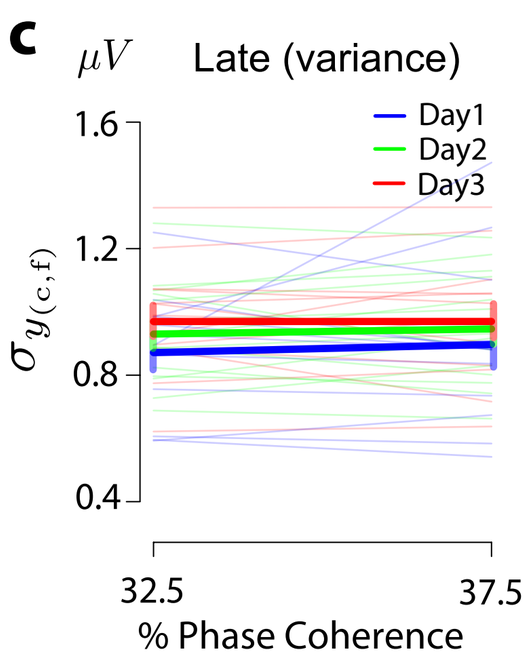
a

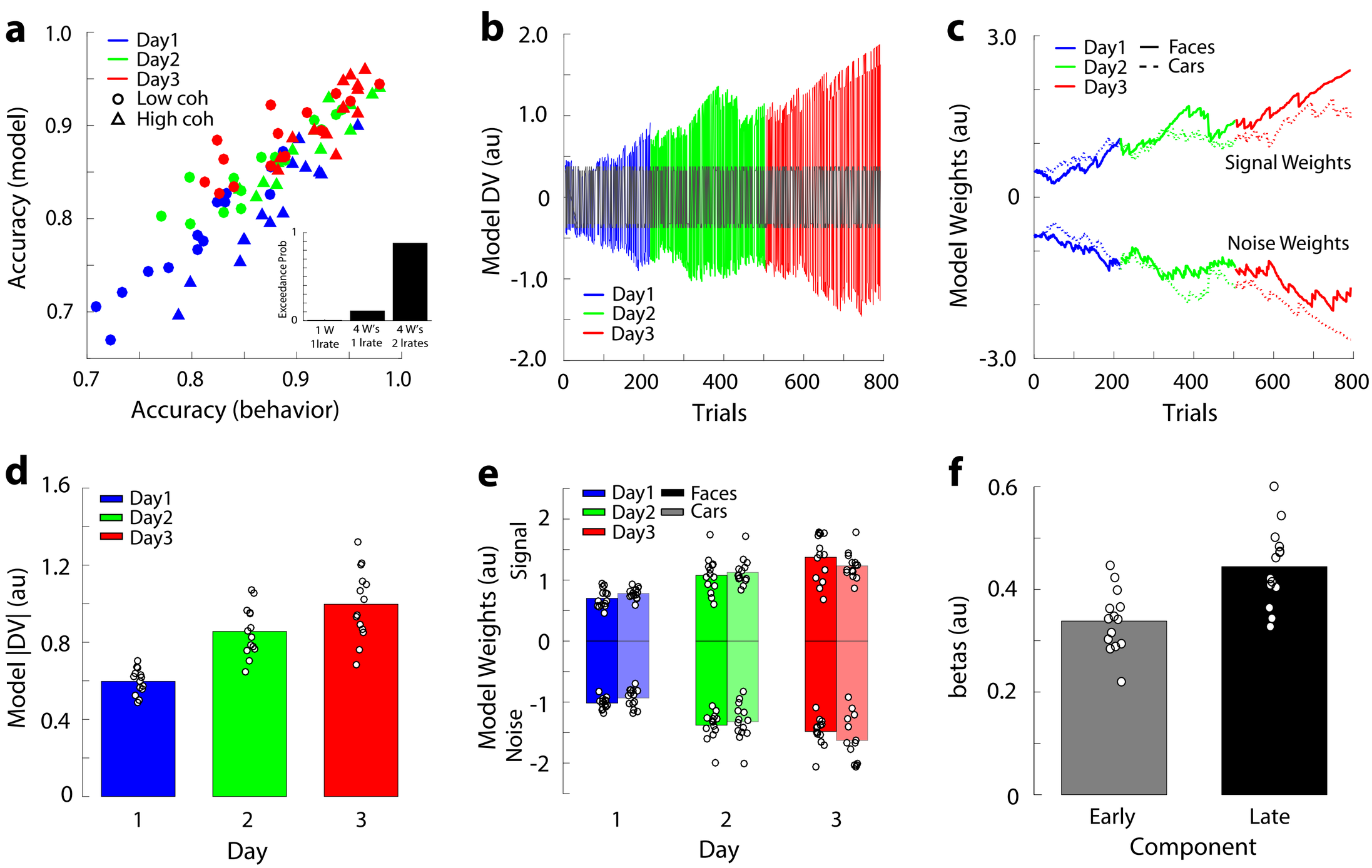

Day 1
Day 2
Day 3
$\bar{y}_\mathrm{f} - \bar{y}_\mathrm{c}$
σ_{y_c}
σ_{y_f}

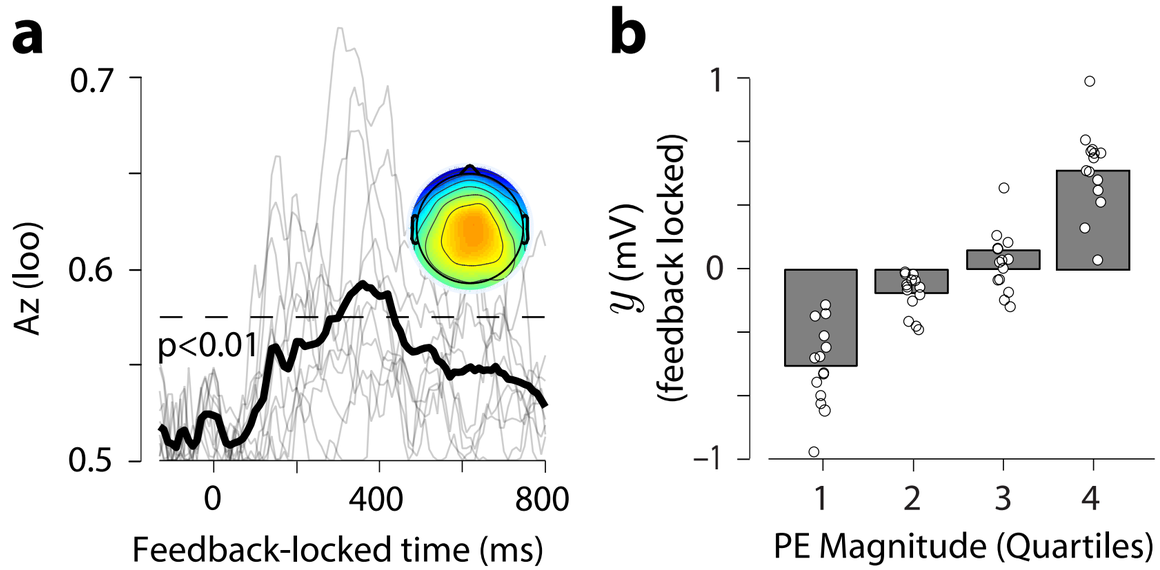
b


μV
Late (mean)
Day1
Day2
Day3
$\bar{y}_\mathrm{f} - \bar{y}_\mathrm{c}$
2.5
2.0
1.5
1.0
0.5
32.5
37.5
% Phase Coherence

c


μV
Late (variance)
Day1
Day2
Day3
$\sigma_{y_{(\mathrm{c,f})}}$
1.6
1.2
0.8
0.4
32.5
37.5
% Phase Coherence


a
Accuracy (model)
Accuracy (behavior)
Day1
Day2
Day3
Low coh
High coh
Exceedance Prob
1 W 1Irate
4 W's 1 Irate
4 W's 2 Irates
b
Model DV (au)
Trials
Day1
Day2
Day3
c
Model Weights (au)
Trials
Day1
Day2
Day3
Faces
Cars
Signal Weights
Noise Weights
d
Model |DV| (au)
Day
Day1
Day2
Day3
e
Model Weights (au)
Signal
Noise
Day
Day1
Day2
Day3
Faces
Cars
f
betas (au)
Component
Early
Late


a
0.7
0.6
0.5
Az (loo)
p<0.01
0
400
800
Feedback-locked time (ms)
b
1
0
−1
y (mV)
(feedback locked)
1
2
3
4
PE Magnitude (Quartiles)